# Phát hiện đối tượng chuyển động

Hầu hết các hệ thống giám sát hình ảnh đều bắt đầu với việc phát hiện đối tượng chuyển động. Phát hiện đối tượng là bước đầu tiên và quan trọng nhất trong hệ thống camera giám sát tự động. Mục tiêu của của việc phát hiện đối tượng chuyển động là xác định vị trí các khu vực đại diện cho các đối tượng di chuyển trong cảnh quay [26] và thực hiện tách chúng ra khỏi nền, mô hình thực hiện được chỉ ra trong hình dưới đây (Hình 1.1). Các quá trình tiếp theo như phân loại đối tượng, theo dõi và phân tích hành vi phụ thuộc rất nhiều vào quá trình này.

**Phát hiện các vùng ảnh nổi  
(foreground detection)**

**Xử lý các vùng ảnh nổi  
(foreground processing)**

**Chuỗi video**

**Các đối tượng chuyển động**

Hình .. Mô hình xử lý trong bài toán phát hiện đối tượng chuyển động

Việc trích rút đối tượng chuyển động khỏi nền phụ thuộc và chịu nhiều tác động bởi các yếu tố như: sự thay đổi cường độ ánh sáng, bóng của đối tượng, sự chuyển động và thay đổi của nền. Wang và Suter [31] cho rằng, một kỹ thuật phát hiện đối tượng tốt cần đạt các yêu cầu sau:

* Phát hiện chính xác hình dạng đối tượng chuyển động;
* Hoạt động tốt trong điều kiện ánh sáng khác nhau (chẳng hạn như bật/tắt ánh sáng, sự chiếu sáng thay đổi dần dần) và sự chuyển động của các đối tượng nền (ví dụ, nếu một đối tượng nền bị di chuyển, thì khi đó đối tượng này không được gán nhãn như là một đối tượng cần theo dõi);
* Linh hoạt với các kịch bản khác nhau (cả trong nhà và ngoài trời);
* Hiệu quả với các mô hình nền khác nhau và hiệu quả trong giai đoạn huấn luyện ngay cả khi các đối tượng tiền cảnh tồn tại trong tất cả các ví dụ huấn luyện;
* Chính xác với cả trường hợp ngụy trang (ví dụ, nếu một đối tượng tiền cảnh có màu tương tự như nền) và độ mở ống kính cận cảnh (nếu di chuyển một đối tượng màu đồng nhất, khi đó một số các điểm ảnh bên trong đối tượng có thể không được phát hiện như là chuyển động);
* Hiệu quả trong tính toán.

Có nhiều phương pháp phát hiện đối tượng chuyển động khác nhau đã được sử dụng, Teddy [26] phân loại thành 3 nhóm chính gồm các kỹ thuật: phương pháp chênh lệch tạm thời, trừ nền, các phương pháp thống kê và các phương pháp sử dụng kỹ thuật luồng quang học.

Sau khi nghiên cứu, khảo sát các công trình nghiên cứu liên quan, luận án phân loại các kỹ thuật phát hiện đối tượng thành ba loại chính, đó là: các phương pháp không sử dụng mô hình nền, có sử dụng mô hình nền và phương pháp tiếp cận tổng hợp.

## Phát hiện đối tượng không sử dụng mô hình nền

Nhóm những kỹ thuật phát hiện đối tượng chuyển động không sử dụng mô hình nền được sử dụng thường xuyên do tính đơn giản và hiệu quả trong tính toán. Ngưỡng hình ảnh, chênh lệch tạm thời và độ lệch không gian – thời gian là các kỹ thuật thường được sử dụng ở lớp này. Các kỹ thuật trong lớp này chỉ sử dụng các khung hình video hiện tại và khung hình trước đó để phát hiện đối tượng chuyển động. Các kỹ thuật trong lớp này chỉ có thể tách các đối tượng chuyển động từ nền tĩnh.

Trong phần này, luận án sẽ trình bày về một số phương pháp hiện có sử dụng ngưỡng hình ảnh, độ lệch thời gian và kỹ thuật dựa trên luồng quang học để tách đối tượng chuyển động và nền.

### Phương pháp sử dụng ngưỡng hình ảnh

Phương pháp sử dụng ngưỡng hình ảnh là phương pháp phát hiện đối tượng đơn giản. Trong kỹ thuật này, người ta cho rằng các đối tượng chuyển động và nền có màu sắc khác nhau. Ritter và Wilson [22] đưa ra phương pháp ngưỡng hình ảnh để phân loại các điểm ảnh là đối tượng hoặc nền. Mỗi điểm ảnh trong khung hình được phân loại là đối tượng hoặc nền. Nếu giá trị màu của một điểm ảnh nằm trong giá trị ngưỡng màu đưa ra, thì sẽ được gán giá trị nhị phân 1 trái lại nó được coi như một điểm ảnh nền và được gán giá trị 0.

Thay vì sử dụng một giá trị ngưỡng toàn cục, thì phương pháp tích lũy, sử dụng ngưỡng cục bộ hoặc ngưỡng động thường được sử dụng [23].

Trong hầu hết các tình huống thực tế, các đối tượng và hình nền được chia sẻ nhiều màu sắc phổ biến, điều này gây ra khó khăn để chọn giá trị ngưỡng tối ưu để tách đối tượng từ nền. Su và Amer [25] tập trung vào hai loại ngưỡng (ước lượng sự phân bố các vùng ảnh thay đổi do vị trí không gian màu) và đề xuất một thuật toán phi tham số để tính toán ngưỡng toàn cục. Phương pháp này cải thiện được việc phát hiện đối tượng. Petrou và Bosdogianni [19] sử dụng biểu đồ màu nhằm tìm ra các giá trị ngưỡng để tách các đối tượng từ nền.

Ngưỡng hình ảnh là một phương pháp hữu ích cho phát hiện đối tượng chuyển động, khi các đối tượng và màu nền là rất khác nhau.

### Phương pháp chênh lệch tạm thời

Phương pháp chênh lệch tạm thời hoặc khác biệt khung hình thực hiện việc phát hiện vùng chuyển động bằng cách tính toán sự khác nhau ở mức điểm ảnh giữa hai hoặc ba khung hình liên tiếp trong một chuỗi các khung hình video [5], chỉ những điểm ảnh có biến động đáng kể (lớn hơn một ngưỡng quy định) được đánh dấu là điểm ảnh của đối tượng chuyển động và điểm ảnh còn lại được phân loại là ảnh nền. Phương pháp này khá tốt đối với khung cảnh động, có nhiều thay đổi song nó lại có nhược điểm trong việc phát hiện ra các điểm ảnh có liên quan đến một số kiểu chuyển động đặc biệt của đối tượng trong cảnh quay.

Phương pháp chênh lệch tạm thời được thực hiện qua công thức

|  |  |
| --- | --- |
|  | (.) |

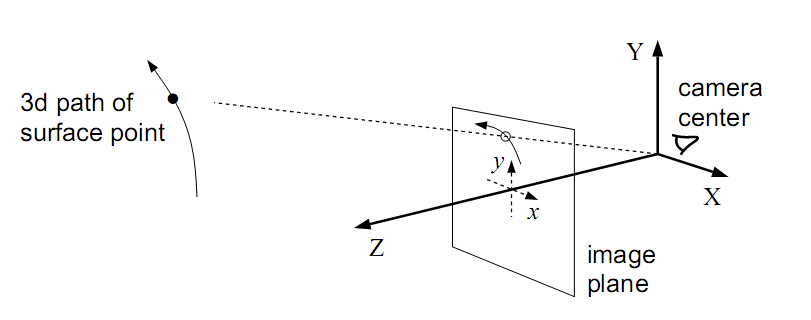
Trong đó là điểm ảnh có tọa độ (x,y) của khung hình tại thời điểm t, là điểm ảnh có tọa độ (x,y) của khung hình tại thời điểm t-1. là ngưỡng được định nghĩa trước đó.

Verbeke và Vincent [30] thay vì tính toán sự khác biệt ngay ở khung hình trước đó, họ tính toán sự khác biệt với mười khung hình đã tích lũy trước đó, kết hợp với phương pháp phân tích thành phần chính để giảm kích thước dữ liệu, kỹ thuật họ đề xuất được đánh giá tốt hơn so với kỹ thuật khác biệt khung hình đơn giản. Tuy nhiên với các điều kiện ánh sáng thay đổi, nhiễu và đặc biệt là hiện tượng dừng tạm thời của đối tượng thì kỹ thuật vẫn chưa giải quyết được. Như các kỹ thuật khác biệt khung hình khác, kỹ thuật này cũng không tách được đối tượng trên nền chuyển động.

### Phương pháp sử dụng luồng quang học

Phương pháp dựa vào luồng quang học (optical flow) thực hiện bằng cách sử dụng các vector có hướng của các đối tượng chuyển động theo thời gian để phát hiện các vùng chuyển động trong một ảnh. Chúng có thể phát hiện chuyển động trong các dãy video ngay cả các video thu được từ camera di chuyển.

Luồng quang học là khái niệm chỉ sự chuyển động tương đối của các điểm trên bề mặt một đối tượng, vật thể nào đó gây ra, dưới góc quan sát của một điểm mốc (mắt, camera…). Sự chuyển động của các vật thể (mà thực tế có thể coi là sự chuyển động của các điểm trên bề mặt của vật thể ấy) trong không gian 3 chiều, khi được chiếu lên một mặt phẳng quan sát 2D được gọi là trường chuyển động (motion field). Nói chung, mục đích của các phương pháp luồng quang học là để xác định (xấp xỉ) trường chuyển động từ một tập các khung hình thay đổi theo thời gian, chúng được sử dụng rộng rãi trong các bài toán phân đoạn đối tượng, phát hiện đối tượng chuyển động, theo vết đối tượng…



Hình .. Luồng quang học

Có nhiều kĩ thuật xác định luồng quang học khác nhau, luận án sẽ tập trung vào phương pháp dựa trên gradient, chi tiết các phương pháp khác có thể tham khảo khảo sát của Raveshiya [21].

Ý tưởng quan trọng của phương pháp tính toán luồng quang học dựa trên gradient là giả định sau:

Bề ngoài của đối tượng không có nhiều thay đổi (về cường độ sáng) khi xét từ frame thứ n sang frame n+1. Nghĩa là:

|  |  |  |
| --- | --- | --- |
|  | (.) |  |

Trong đó là hàm trả về cường độ sáng của điểm ảnh tại thời điểm t (khung hình thứ t). là tọa độ của điểm ảnh trên bề mặt (2D), là vector vận tốc, thể hiện sự thay đổi vị trí của điểm ảnh từ frame thứ t sang frame t+1. Mặc dù có vẻ không thực tế, nhưng nhiều công trình cho thấy giả sử (1.2) làm việc khá tốt trong thực tế.

Sử dụng khai triển Taylor, ta xấp xỉ:

|  |  |  |
| --- | --- | --- |
|  | (.) |  |

với , và là đạo hàm tức thời của I tại thời điểm t.

Bỏ qua các phần tử có bậc đạo hàm cao hơn trong (1.3), so sánh (1.2) và (1.3), đồng thời xấp xỉ bằng , ta nhận được phương trình sau, gọi là phương trình ràng buộc gradient (gọi là phương trình luồng quang học):

|  |  |  |
| --- | --- | --- |
|  | (.) |  |

Phương trình (1.4) còn có thể đạt được bằng cách xuất phát ngay từ giả thiết ban đầu là giá trị cường độ sáng của điểm ảnh không thay đổi từ frame t sang t+1, như vậy:

|  |  |  |
| --- | --- | --- |
|  | (.) |  |

Phương trình (1.4) là phương trình cơ bản của luồng quang học, trong đó  chính là luồng cần tìm. Tuy nhiên phương trình này có dạng  với 2 ẩn là và nên không giải được. Ta có nhận xét rằng phương trình (4), với ẩn , là phương trình đường thẳng trong 2D với vector pháp tuyến là .

Mọi phương pháp ước lượng luồng quang học đều xoay quanh việc cố gắng tìm thêm một phương trình nữa để tính .

Để tìm thêm một ràng buộc nữa cho , ta có thể dùng đến các điểm ảnh lân cận với điểm ảnh đang xét, với giả sử là các điểm ảnh lân cận này cũng nằm trên vật thể và chuyển động với cùng vận tốc của điểm ảnh đang xét. Trong thực tế, các điểm ảnh lân cận có thể không chuyển động cùng vận tốc với điểm ảnh đang xét, do đó ta tìm vector vận tốc mà làm cực tiểu bình phương lỗi:

|  |  |  |
| --- | --- | --- |
|  | (.) |  |

Phương pháp này gọi là ước lượng bình phương tối thiểu (Least-squares estimation), trong đó xác định “độ rộng” mà hàm ước lượng này bao phủ. Thông thường người ta chọn hàm Gaussian. Vận tốc cần tìm là giá trị làm cực tiểu .

Người ta chứng minh được đạt cực tiểu tại điểm đạo hàm riêng của nó bằng zero, nghĩa là:

|  |  |
| --- | --- |
|  | (.) |
|  | (.) |

Hai phương trình này có thể viết ở dạng ma trận:

|  |  |
| --- | --- |
|  | (.) |

với:

|  |  |
| --- | --- |
|  | (.) |

Nghiệm của (1.9) là 

Các phương pháp luồng quang học rất phù hợp với tình huống camera di chuyển, tuy nhiên các phương pháp này luôn đòi hỏi sự tính toán khá phức tạp và khá nhạy với nhiễu.

## Phát hiện đối tượng chuyển động sử dụng mô hình nền

Các kỹ thuật sử dụng mô hình nền (hay trừ nền), có khả năng phân đoạn hình ảnh thành các đối tượng chuyển động và nền trong một số điều kiện phức tạp như: nền chuyển động, sự thay đổi độ sáng và cũng có thể giải quyết được cả khi đối tượng dừng tạm thời tại một số khung hình.

Mặc dù có nhiều kỹ thuật trừ nền khác nhau, nhưng phần lớn các kỹ thuật trừ nền đều có một đặc điểm chung, đó là giả định rằng dãy video quan sát các đối tượng được thực hiện trên một nền tĩnh B và các đối tượng chuyển động có mầu sắc khác với mầu của nền B. Kỹ thuật này đơn giản chỉ tính sự khác biệt giữa ảnh hiện tại và ảnh nền nhằm tìm ra vùng các điểm ảnh khác nhau . Các điểm ảnh khác nhau sẽ được hiểu như là các điểm ảnh nổi lên trên ảnh nền. Sau khi tìm ra các vùng ảnh nổi, các vùng này sẽ được xử lý để lọc đi các nhiễu, các vết không phù hợp bằng một số thuật toán lọc nhiễu khác nhau. Ảnh nền thông thường được chọn là khung hình đầu tiên của video không có các đối tượng tiền cảnh.

|  |  |
| --- | --- |
|  | (.) |

Trong đó là ngưỡng được xác định trước. Nếu giá trị tuyệt đối lớn hơn hoặc bằng điểm ảnh đó được xác định là điểm ảnh nổi (foreground), ngược lại điểm ảnh đó là nền. Sự khác biệt chính của các phương pháp trừ nền là cách thức mô hình hóa B và đơn vị đo khoảng cách d được sử dụng [1]. Yannick Benezeth đã có khảo sát chi tiết về các phương pháp trừ nền [1].

Cách đơn giản nhất để mô hình hóa nền B là thông qua một ảnh xám hoặc màu không chứa đối tượng chuyển động. Hình ảnh này có thể được chụp trong trường hợp không có chuyển động hoặc có thể ước lượng được thông qua một bộ lọc trung bình theo thời gian [4], [11], [33].

Ghidary cùng đồng nghiệp [8] sử dụng khung hình trước I(x,y)t-1 như là ảnh nền B, tuy nhiên, với cấu hình này việc phát hiện chuyển động trở thành quá trình phát hiện sự thay đổi giữa các khung hình, điều này khá hiệu quả đối với việc thay đổi ánh sáng, nhưng do vấn đề về độ mở ống kính nên chỉ một phần của đối tượng chuyển động được phát hiện.

Ảnh nền B cũng được tạo bằng cách tính ảnh trung bình của khung hình hiện tại với n khung hình trước đó [3], [17]. Xét về độ phức tạp tính toán, thì phương pháp trừ nền tính toán nhanh nhưng đòi hỏi nhiều bộ nhớ hơn để lưu trữ n khung hình trước [20].

Một trong các phương pháp phổ biến khác đó là tiếp cận dựa vào xác suất, Stauffer và Grimson [24] sử dụng mô hình hỗn hợp Gaussian (GMM), kỹ thuật đề xuất có khả năng thích ứng tốt với sự thay đổi chiếu sáng dần dần cũng như sự dịch chuyển của nền. Tuy nhiên, kỹ thuật của họ không đáp ứng được với điều kiện cường độ ánh sáng thay đổi đột ngột.

Thome và Miguet [27] sử dụng kết hợp giữa GMM với kỹ thuật loại bỏ bóng, kỹ thuật họ đề cho kết quả tốt hơn so với kết quả khẳng định trong [24].

Dickinson và đồng nghiệp [6] thực hiện mô hình hóa nền bằng mô hình GMM thích nghi với màu sắc và không gian, kỹ thuật đề xuất cho kết quả tốt hơn so với GMM truyền thống.

Elgammal [7] giới thiệu một phương pháp ước lượng dựa trên mật độ hạt nhân và cho thấy phương pháp đề xuất có hiệu quả trong việc xử lý các tình huống mà nền có chứa di chuyển nhỏ và lặp đi lặp lại như cành cây và bụi cây. Tuy nhiên chi phí để tính toán mật độ hạt nhân tại mỗi điểm ảnh là rất cao, các tác giả đưa ra một số bảng tra cứu tính trước để giảm bớt gánh nặng tính toán của thuật toán

Tương tự như vậy, Han và đồng nghiệp [10] mô hình hóa nền bằng một xấp xỉ mật độ tuần tự và mỗi bước thời gian, mật độ được ước tính và thành phần Gaussian được gán cho mỗi mô hình.

Pic và cộng sự [18] sử dụng một kỹ thuật thích ứng cho việc ước lượng nền trên cơ sở học. Tỷ lệ học được tính toán sau mỗi khung cho mỗi điểm ảnh, do vậy mà chi phí cho tính toán khá cao. Kỹ thuật không đáp ứng được với những thay đổi nhanh của đối tượng trong tiền cảnh và nền, đồng thời nó cũng nhạy hơn với sự thay đổi của cường độ ánh sáng.

Kim và đồng nghiệp [15] thực hiện phân đoạn tiền cảnh và nền bằng cách sử dụng phương pháp codebook. Kỹ thuật này cho thấy hiệu suất phát hiện đối tượng tốt và nó cũng hiệu quả hơn cho khi xử lý vấn đề bóng và sự biến đổi cường độ ánh sáng.

Phương pháp VuMeter [9] đề xuất là một mô hình phi tham số. Đây là một cách tiếp cận xác suất để xác định mô hình nền hình ảnh sử dụng ước lượng của hàm phân bố xác suất. Một điểm ảnh có thể có hai trạng thái, hoặc là nền hoặc tiền cảnh. Mỗi điểm ảnh nền được cập nhật bằng cách sử dụng tỷ lệ học cố định.

Zhong và Sclaroff [32] đề xuất một thuật toán thực hiện mô hình hóa một cách rõ ràng cấu trúc nền động thông qua một bộ lọc Kalman, thuật toán được sử dụng để ước tính sự xuất hiện nội tại của các kết cấu động. Các vùng đối tượng tiền cảnh thu được bằng cách đặt ngưỡng cho hàm trọng số sử dụng trong bộ lọc Kalman.

Các đối tượng tĩnh sẽ chuyển thành nền nếu chúng dừng lại một vài giây và kỹ thuật mô hình nền chấp nhận màu đối tượng tĩnh trong nền. Sự thay đổi đột ngột của cường độ ánh sáng làm cho mô hình nền không ổn định. Phương pháp được đề xuất bởi [15] thực hiện tốt hơn trong những tình huống và cách tiếp cận khác có thể vượt qua các vấn đề ở trên bằng cách sử dụng phương pháp kết hợp.

## Tiếp cận kết hợp

Các kỹ thuật trong lớp này kết hợp các kỹ thuật sử dụng mô hình hóa nền với các thuật toán khác. Ví dụ, kết hợp chênh lệch không gian – thời gian với kỹ thuật mô hình nền. Tian [29] cho rằng các đối tượng tiền cảnh bị hấp thụ ở các mức khác nhau tại các điểm ảnh khác nhau, gây ra sự phân mảnh đối tượng. Vấn đề phân mảnh cũng phát sinh khi các đối tượng tiền cảnh chồng chéo lên nhau với các đối tượng nền có màu sắc tương tự. Các loại lỗi là khó tránh khỏi theo giả định của một mô hình điểm ảnh độc lập. Hình ảnh cảnh được tạo ra bởi một tập các đối tượng rời rạc (bao gồm nền và tiền cảnh) bởi vậy giá trị các điểm ảnh được tạo ra bởi cùng một đối tượng thể hiện một sự gắn kết không gian, màu sắc, và thời gian chặt chẽ.

Javed và cộng sự [14] sử dụng GMM và cường độ gradient để loại bỏ bóng và bù cho sự thay đổi của cường độ ánh sáng.

Li và cộng sự, [16] đề xuất sử dụng bộ khung Bayesian kết hợp đặc tính quang phổ, không gian và thời gian để mô tả sự xuất hiện nền tại mỗi điểm ảnh. Họ cho rằng phương pháp của họ có thể xử lý cả hai trường hợp nền tĩnh và động. Thuật toán của họ giảm đáng kể nếu các đối tượng tiền cảnh liên tục xuất hiện trong cảnh.

Heikkilä và Pietikäinen [12] mô hình mỗi điểm ảnh bởi một nhóm các mẫu nhị phân cục bộ (Local Binary Pattern - LBP), biểu đồ được tính toán trên một vùng hình tròn xung quanh điểm ảnh. Phương pháp này xử tốt tốt với nền động vì nó tập hợp thông tin trên một khu vực chứ không phải là một điểm ảnh. Tuy nhiên, có một nhược điểm, đó là sự không chính xác của các thông tin hình dạng của kết quả phân đoạn do dựa trên việc phân vùng. Một vấn đề khác của phương pháp này là tỷ lệ học chậm.

Izadi và Saeedi [13] kết hợp gradient không gian với GMM để phát hiện các đối tượng. Họ tách các đối tượng từ nền và cũng loại bỏ bóng bằng cách sử dụng bộ lọc và các phép toán hình thái.

Cong và đồng nghiệp [2] phát hiện các đối tượng chuyển động bằng cách kết hợp kỹ thuật mô hình nền GMM với độ lệch thời gian các khung hình kế tiếp.

Tian và Men [28] thay đổi phương pháp của [12] bằng cách giới thiệu một biểu đồ không gian trọng số LBP (Spatially Weighted LBP Histogram - SWLH) như một vector đặc trưng. SWLH cải thiện độ chính xác thông tin hình dạng tốt hơn so với [12].

Giống như tất cả các lĩnh vực khác, kết hợp nhiều thuật toán sẽ cải thiện được hiệu suất của việc phát hiện đối tượng nhưng làm cho việc tính toán tốn kém hơn.

TÀI LIỆU THAM KHẢO

1. Benezeth Y., Jodoin P.-M., Emile B., Laurent H., and Rosenberge C. (2010). "Comparative study of background subtraction algorithms", *J. Electron. Imaging, 19*(3).

2. Cong D. N. T., Khoudour L., Achard C., and Phothisane P. (2009). "People re-identification by means of a camera network using a graph-based approach", *Machine Vision and Applications, 90*, pp: 2362–2374.

3. Cucchiara R., Grana C., Piccardi M., and Prati A. (2003). "Detecting moving objects, ghosts and shadows in video streams", *PAMI, Vol 25*(10), pp: 1337–1342.

4. Cucchiara R., Grana C., Prati A., and Vezzani R. (2005). "Probabilistic posture classification for human-behavior analysis", *Transactions on Systems, Man and Cybernetics, Part A*(35), pp: 42–54.

5. Dedeo˘glu Y. g. (2004). "Moving object detection, tracking and classification for smart video surveillance", *A thesis submitted to the department of computer engineering and the institute of engineering and science of bilkent university in partial fulfillment of the requirements for the degree of master of science*.

6. Dickinson P., Hunter A., and Appiah K. (2003). "Segmenting foreground objects from a dynamic textured background via a robust kalman filter", *IEEE International Conference on Computer Vision*, pp: 44–50.

7. Elgammal A., Harwood D., and Davis L. (2000). "Non-parametric model for background subtraction", *In 6th European Conference on Computer Vision, ECCV00*, pp: 751–767.

8. Ghidary S., Nakata Y., Takamori T., and Hattori M. (2000). "Human detection and localization at indoor environment by homerobot", *International Conference on Systems, Man, and Cybernetics, Vol 2*, pp: 1360–1365.

9. Goyat Y., Chateau T., Malaterre L., and Trassoudaine L. (2006). "Vehicle trajectories evaluation by static video sensors", *In 9th IEEE International Conference on Intelligent Transportation Systems*, pp: 864–869.

10. Han B., Comaniciu D., and Davis L. (2008). "Sequence kernel density approximation through mode propagation: Application to background modeling", *Proceedings of the IEEE Computer Vision and Pattern Recognition, PP(99)*, pp: 1186–1197.

11. Heikkila J., and Silven O. (2004). "A real-time system for monitoring of cyclists and pedestrians", *Workshop on Visual Surveillance*, pp: 74–81.

12. Heikkilä M., and Pietikäinen M. (2006). "A texture based method for modeling the background and detecting moving objects", *Transaction on Pattern Analysis and Machine Intelligence, 28*(4), pp: 657–662.

13. Izadi M., and Saeedi P. (2008). "Robust region-based background subtraction and shadow removing using color and gradient information", *In 19th International Conference on Pattern Recognition ICPR*, pp: 1–5.

14. Javed O., Shafique K., and Shah M. (2002). "A hierarchical approach to robust background subtraction using color and gradient information", *Motion and Video Computing*, pp: 22–27.

15. Kim K., Thanarat T., Chalidabbhognse H., Harwood D., and Davis L. (2005). "Real time foreground-background segmentation using codebook model", *Real-Time Imaging, 11*(2), pp: 172–185.

16. Li L., Huang W., Gu I. Y. H., and Tian Q. (2004). "Statistical modeling of complex backgrounds for foreground object detection", *IEEE Transaction on Image Processing, 13*(11), pp: 1459–1472.

17. Lo B. P. L., and Velastin S. A. (2001). "Automatic congestion detection system for underground platforms", *In International Symposium on Intelligent Multimedia, Video and Speech Processing*, pp: 158–161.

18. M. Pic, Berthouze L., and Kurita T. (2004). "Adaptive background estimation: Computing a pixel-wise learning rate from local confidence and global correlation", *IEICE Transaction Information and System, E87-D(1)*, pp: 50–57.

19. Petrou M., and Bosdogianni P. (2010), *Image Processing: The Fundamental*, John Wiley and Sons Inc, New York, USA.

20. Piccardi M. (2004). "Background subtraction techniques: A review", *In IEEE International Conference on Systems, Man and Cybernetics*.

21. Raveshiya H., and Borisagar V. (2012). "Motion Estimation Using Optical Flow Concepts", *Int.J.Computer Technology & Applications, 3*(2), pp: 696-700.

22. Ritter G. X., and Wilson J. N. (2000). *Handbook of Computer Vision Algorithms in Image Algebra* [2nd edition version].

23. Shapiro L. G., and Stockman G. C. (2002). *Computer Vision*.

24. Stauffer C., and Grimson W. E. L. (2000). "Learning patterns of activity using real-time tracking", *IEEE Transactions on Pattern Analysis and Machine Intelligence, 22*(8), pp: 747–757.

25. Su C., and Amer A. (2006). "A real-time adaptive thresholding for video change detection", *In International Conference on Image Processing, ICIP06*, pp: 157–160.

26. Teddy K. (2011). A Survey on Behavior Analysis in Video Surveillance Applications. *Video Surveillance* (16, pp. 279 - 294): InTech.

27. Thome N., and Miguet S. (2005). "A robust appearance model for tracking human motion", *Advanced Video and Signal Based Surveillance (AVSS)*, pp: 528–533.

28. Tian G. D., and Men A. D. (2009). "An improved texture-based method for background subtraction using local binary patterns", *In CISP09*, pp: 1–4.

29. Tian Y., Lu M., and Hampapur A. (2005). "Robust and efficient foreground analysis for realtime video surveillance", *In IEEE Conference on Computer Vision and Pattern Recognition CVPR0*, pp: 1182–1187.

30. Verbeke N., and Vincent N. (2007). "A PCA-based technique to detect moving objects", *In SCIA*, pp: 641–650.

31. Wang H., and Suter D. (2007). "A consensus-based method for tracking: Modelling background scenario and foreground appearance", *Pattern Recognition, Vol 40*(3), pp: 1091 – 1105.

32. Zhong J., and Sclaroff S. (2003). "Segmenting foreground objects from a dynamic textured background via a robust kalman filter", *In IEEE International Conference on Computer Vision (ICCV)*, pp: 44–50.

33. Zhou Q., and Aggarwal J. (2001). "Tracking and classifying moving objects from video", *Performance Evaluation of Tracking Systems Workshop*.